Expression in *Escherichia coli* of fragments of the coiled-coil rod domain of rabbit myosin: influence of different regions of the molecule on aggregation and paracrystal formation

SIMON J. ATKINSON* and MURRAY STEWART†

MRC Laboratory of Molecular Biology, Hills Rd, Cambridge CB2 2QH, UK

* Present address: Department of Cell Biology and Anatomy, Johns Hopkins Medical School, Baltimore, MD 21205, USA
† Author for correspondence

Summary

We have expressed in *Escherichia coli* a cDNA clone corresponding broadly to rabbit light meromyosin (LMM) together with a number of modified polypeptides and have used this material to investigate the role of different aspects of molecular structure on the solubility properties of LMM. The expressed material was characterized biochemically and structurally to ensure that it retained the coiled-coil conformation of the native molecule. Full-length recombinant LMM retained the general solubility properties of myosin and, although soluble at high ionic strength, precipitated when the ionic strength was reduced below 0.3 M. Constructs in which the 'skip' residues (that disrupt the coiled-coil heptad repeat) were deleted had solubility properties indistinguishable from the wild type, which indicated that the skip residues did not play a major role in determining the molecular interactions involved in assembly. Deletions from the N terminus of LMM did not alter the solubility properties of the expressed material, but deletion of 92 residues from the C terminus caused a large increase in solubility at low ionic strength, indicating that a determinant important for interaction between LMM molecules was located in this region. The failure of deletions from the molecule's N terminus to alter its solubility radically suggested that the periodic variation of charge along the myosin rod may not be as important as proposed for determining the strength of binding between molecules and thus the solubility of myosin.

Key words: myosin, coiled-coil, crystals, paracrystals, molecular interactions, expression.

Introduction

Myosin is an important motor protein in both muscle and non-muscle cells (reviewed by Harrington and Rogers, 1984; Warwick and Spudich, 1987). The function of myosin depends on its ability to polymerize into macromolecular assemblies that enable individual molecules to join forces and produce a relative sliding of actin-containing filaments. In striated muscle, myosin assembles into long, bipolar thick filaments, which have been characterized extensively (Squire, 1981; Vibert and Craig, 1983; Crowther *et al.* 1985; Stewart *et al.* 1985*b*; Cantino and Squire, 1986; Stewart and Kensler, 1986; Kensler and Stewart, 1989). Myosin also aggregates in smooth- and non-muscle systems, but in these cases the structure of the aggregates has not been determined with the same precision as for skeletal muscle. *In vitro*, myosin aggregates to form assemblies similar to those seen *in vivo*, and so this is essentially a self-assembly process (Huxley, 1963). Consequently, although other factors may contribute to the precise arrangement *in vivo* (and in particular may control the length of the filament; see Atkinson and Stewart, 1991), the structure of the thick filament is determined largely by the properties of the myosin molecules and the manner in which they interact with one another.

Myosin molecules have two globular heads (that interact with actin to form the cross bridges that bring about the relative sliding of the thick and thin filaments) and a long rod-like tail that is formed by two alpha-helices wrapping around one another to form a coiled-coil (reviewed by McLachlan, 1984; Warwick and Spudich, 1987). The head and tail domains can be separated by proteolysis (Lowey *et al.* 1969) and the tail can be further cleaved to give an N-terminal fragment, termed subfragment-2 (S-2), and a C-terminal fragment, light meromyosin (LMM). LMM retains the solubility properties of the intact myosin molecule and, while soluble at high ionic strength, aggregates under physiological conditions.

Although there are differences in detail, there are several unifying themes in the manner in which myosin heads are arranged on the surface of thick filaments from a wide variety of species. In vertebrate skeletal muscle, for example, the heads are arranged on a right-handed three-stranded approximately helical lattice, with a helical repeat of 43 nm, and a 14.3 nm axial translation between successive crossbridge levels (Huxley and Brown, 1967; Kensler and Stewart, 1983; Ip and Heuser, 1983; Cantino and Squire, 1986; Stewart and Kensler, 1986). Invertebrate muscle thick filaments have different numbers of helical strands (4 in the case of arthropod muscles (Stewart *et al.* 1981), 7 in the case of scallop striated

muscle (Vibert and Craig, 1983)), but the helical repeat (43 nm) and the axial translation between crossbridge levels (14.3 nm) are the same. These parameters probably reflect the arrangement of the myosin rods in the core of the thick filament, but this arrangement has not been established in detail. These 43 nm and 14.3 nm repeats are also found in paracrystals formed when the ionic strength of LMM solutions is reduced to near physiological levels. The paracrystals have distinct axial banding patterns that are derived from the axial spacing of LMM molecules within them and almost invariably these patterns are based on multiples of 14.3 nm and commonly of 43 nm (Kendrick-Jones *et al.* 1971; Katsura and Noda, 1973; Safer and Pepe, 1978; Yagi and Offer, 1981; Quinlan and Stewart, 1987; Ward and Bennett, 1989). These axial staggers between molecules in paracrystals are thought to reflect the preferred axial staggers between myosin rods in thick filaments and so give a clue to the molecular interactions important for filament assembly. Myosin becomes increasingly soluble as the ionic strength and pH are raised (Katsura and Noda, 1973). These data, together with ultracentrifugation studies of myosin association (Godfrey and Harrington, 1970*a*,*b*), indicate that interactions involving ions (such as the formation of salt bridges and the binding of small anions and cations) are important for assembly.

All myosin rod sequences share common sequence motifs (reviewed by McLachlan, 1984; Warrick and Spudich, 1987), based on the heptad repeat characteristic of sequences that form an alpha-helical coiled-coil (Crick, 1953; McLachlan and Stewart, 1975). Most prominent amongst these is a periodicity in the sign of the charged residues with a repeat of 28 residues (Parry, 1981; McLachlan and Karn, 1982, 1983), so that overall there is an alternation of positively and negatively charged zones along the molecule. This pattern suggests that electrostatic interactions between molecules could provide the energy necessary for myosin assembly if the molecules were staggered by an odd multiple of 28/2 residues (Parry, 1981; McLachlan and Karn, 1982, 1983). In this context, it is significant that the two staggers seen in thick filaments and also in paracrystals both satisfy this condition: staggers of 98 residues (7×28/2) and 294 residues (21×28/2) correspond to axial translations of 14.3 and 43 nm, respectively. However, a very large number of other staggers also satisfy this condition and so simple complementation of charged residues cannot be the complete explanation for the molecular interaction geometry observed in thick filaments. It may be that the coiled-coil pitch of 14.3 nm, established for crystals of the long S-2 fragment (Quinlan and Stewart, 1987), plays a role in constraining staggers to multiples of 14.3 nm.

At four positions along the 1000-residue long myosin rod, the regular heptad periodicity of hydrophobic residues is interrupted by the insertion of an additional 'skip' residue that may disrupt the local structure of the coiled-coil (Parry, 1981; McLachlan and Karn, 1983; McLachlan, 1984). Such a local disruption may be consistent with bends seen in isolated myosin molecules (Offer, 1990) and may have a role in the precise packing of molecules into thick filaments.

We describe here the expression of a cDNA clone corresponding to most of the LMM region of rabbit skeletal muscle myosin. Fragments with regions deleted from the N and C termini of the recombinant LMM were also expressed to define regions of the rod important for assembly. These expressed proteins were characterized in detail and their solubility properties and paracrystal structure investigated. These results indicate that, although the 28-residue charge repeat may be important for determining the relative axial arrangement of molecules in the thick filament, the ability to assemble is conferred by interactions that specifically require the C-terminal region of LMM.

Materials and methods

Materials

All reagents were of analytical grade. Radiochemicals were obtained from Amersham plc (Amersham, Bucks). Oligodeoxynucleotides were synthesized by Terry Smith and Jan Fogg of this laboratory using an Applied Biosystems (Warrington, UK) 380B automated DNA synthesizer by the phosphate triester method (Gait, 1984). Restriction endonucleases, polynucleotide kinase and bacteriophage T4 DNA ligase were from New England Biolabs (Beverley, MA). Klenow fragment of DNA polymerase, calf intestinal alkaline phosphatase, and bovine pancreatic RNAase were from Boehringer Mannheim (Mannheim, FRG). Hen egg white lysozyme and chymotrypsin were from Sigma Chemical Company (St Louis, MO). Expression vector pIN II A2 (Nakamura and Inouye, 1982) was a generous gift from Dr Masayori Inouye (SUNY, Stony Brook, NY).

General microbiological manipulations

Routine manipulations of bacteria and nucleic acids were essentially as described by Maniatis *et al.* (1982). *Escherichia coli* K-12 strain JM101 (Messing, 1983) was used for routine cloning manipulations using M13, pUC and pIN vector systems. Expression was carried out in strain HB2151 and mutagenesis in strain BMH 71–18 (Carter *et al.* 1985). For liquid culture of JM101, HB 2151 and BMH 71–18 strains of *E. coli*, 2×TY medium (Messing, 1983) was used. Strains carrying M13 and plasmids were maintained on TYE plates (Nagai and Thogersen, 1987). For selection of plasmids carrying the AmpR gene, ampicillin (100 $\mu g\,ml^{-1}$) was added to plates and media. DNA sequencing was carried out by the dideoxynucleotide chain termination method of Sanger *et al.* (1977).

Biochemical methods

Concentrations of expressed proteins were usually determined by the Bradford (1976) method, using either BSA (Sigma pentax fraction V) or lyophilized LMM (from rabbit muscle) as standard. Myosin and LMM were determined spectrophotometrically, using extinction coefficients of $0.53\,ml\,mg^{-1}\,cm^{-1}$ for myosin, $0.22\,ml\,mg^{-1}\,cm^{-1}$ for rod, and $0.3\,ml\,mg^{-1}\,cm^{-1}$ for LMM at 280 nm (Margossian and Lowey, 1982). SDS–PAGE was as described (Laemmli, 1970) using microslab gradient gels. Protein sequencing was carried out by sequential Edman degradation using an Applied Biosystems gas-phase sequenator by Mr Ross Jakes of this laboratory. Western blotting was as described (Quinlan *et al.* 1989) using a broad-spectrum myosin monoclonal antibody (Parke *et al.* 1986) obtained from Amersham plc.

Isolation and sequencing of a rabbit myosin cDNA clone

A 1.4 kb cDNA clone encoding part of LMM, which had been isolated serendipitously whilst screening a rabbit muscle cDNA library (constructed in lambda gt10) for glycogenin (Da Cruz e Silva and Cohen, 1987), was a generous gift from Dr Patricia Cohen (University of Dundee, Scotland). The insert was labelled with [α-^{32}P]dCTP using a random oligo kit (Pharmacia) and used to re-probe the library for longer inserts by standard methods (Maniatis *et al.* 1982). The phage giving the strongest hybridization signals were plaque purified and the *Eco*RI inserts subcloned to pUC18. The longest of these (1.7 kb) was sequenced on both strands throughout its length (Fig. 1), using overlapping restriction fragments (Fig. 2) subcloned into M13mp18 and mp19. The sequence has been submitted to the EMBL/Genbank database with accession number X59602.

```
 E  F  G  L  E  E  A  K  K  K  L  A  Q  R  L  Q  D  S  E  E  Q  V  E  A  V  N  A  K  C  A
GAATTCGGGCTGGAGGAGGCCAAGAAAAAACTAGCTCAGCGCCTGCAGGATTCCGAGGAACAGGTTGAGGCTGTGAATGCTAAGTGCGCG
        10        20        30        40        50        60        70        80        90
 S  L  E  K  T  K  Q  R  L  Q  G  E  V  E  D  L  M  V  D  V  E  R  A  N  S  L  A  A  A  L
TCCCTGGAGAAGACCAAGCAGAGGCTGCAAGGGGAGGTGGAGGACCTGATGGTGGACGTGGAGAGGGCCAACTCCCTGGCCGCCGCTCTG
       100       110       120       130       140       150       160       170       180
 D  K  K  Q  R  N  F  D  K  V  L  A  E  W  K  T  K  C  E  E  S  Q  A  E  L  E  A  S  L  K
GACAAGAAGCAGAGGAACTTCGACAAGGTGCTGGCCGAATGGAAGACCAAGTGTGAGGAGAGCCAGGCGGAGCTGGAGGCGTCTCTCAAG
       190       200       210       220       230       240       250       260       270
 E  S  R  T  L  S  T  E  L  F  K  L  K  N  A  Y  E  E  A  L  D  Q  L  E  T  V  K  R  E  N
GAGTCCCGCACCTTGAGCACTGAGCTTTTCAAGCTGAAAAATGCCTACGAGGAGGCCTTAGATCAACTTGAAACTGTGAAACGGGAGAAT
       280       290       300       310       320       330       340       350       360
 K  N  L  E  Q  E  I  A  D  L  T  E  Q  I  A  E  N  G  K  T  I  H  E  L  E  K  S  R  K  Q
AAGAACTTAGAACAAGAGATAGCGGATCTCACAGAGCAAATTGCGGAAAATGGTAAAACCATCCACGAACTGGAGAAATCCAGGAAGCAG
       370       380       390       400       410       420       430       440       450
 M  E  L  E  K  A  D  I  Q  L  A  L  E  E  A  E  A  A  L  E  H  E  E  A  K  I  L  R  V  Q
ATGGAGCTGGAGAAGGCTGACATCCAGCTGGCTCTCGAGGAAGCAGAGGCTGCTCTCGAACACGAAGAAGCCAAGATCCTCCGGGTCCAG
       460       470       480       490       500       510       520       530       540
 L  E  L  T  Q  V  K  A  E  I  D  R  K  I  A  E  K  D  E  E  M  E  Q  L  K  R  N  Y  Q  R
CTCGAGCTGACGCAGGTGAAAGCCGAGATCGACAGGAAGATTGCCGAGAAGGACGAGGAGATGGAGCAGCTGAAGAGGAACTACCAGCGG
       550       560       570       580       590       600       610       620       630
 T  V  E  T  M  Q  S  A  L  D  A  E  V  R  S  R  N  E  A  I  R  I  K  K  K  M  E  G  D  L
ACGGTGGAGACCATGCAGAGCGCCCTGGACGCCGAGGTGCGCAGCAGGAACGAAGCCATCCGCATCAAGAAGAAGATGGAGGGGGACCTG
       640       650       660       670       680       690       700       710       720
 N  E  I  E  I  Q  L  S  H  A  N  R  Q  A  A  E  T  L  K  H  L  R  S  V  Q  G  Q  L  K  D
AACGAGATCGAGATCCAGTTGAGCCACGCCAACCGCCAGGCCGCCGAGACCCTCAAGCACCTCCGGAGCGTCCAGGGACAGCTGAAGGAT
       730       740       750       760       770       780       790       800       810
 T  Q  L  H  L  D  D  A  L  R  G  Q  E  D  L  K  E  Q  L  A  L  V  E  R  R  A  N  L  L  Q
ACCCAGCTGCACCTGGACGATGCTCTCCGCGGCCAGGAGGACCTGAAGGAGCAGCTGGCCCTGGTGGAGCGCAGGGCCAACCTGCTGCAG
       820       830       840       850       860       870       880       890       900
 A  E  V  E  E  L  R  A  T  L  E  Q  T  E  R  A  R  K  L  A  E  Q  E  L  L  D  A  N  E  R
GCTGAGGTGGAGGAGCTGCGGGCCACGCTGGAGCAGACGGAGAGGGCCCGGAAACTGGCTGAGCAGGAGCTCCTGGATGCCAATGAGCGG
       910       920       930       940       950       960       970       980       990
 V  Q  L  L  H  T  Q  N  T  S  L  I  H  T  K  K  K  L  E  T  D  L  T  Q  L  Q  S  E  V  E
GTGCAGCTGCTGCACACCCAGAACACCAGCCTCATCCACACCAAGAAGAAGCTGGAGACGGACCTCACCCAGCTGCAGAGCGAGGTGGAG
      1000      1010      1020      1030      1040      1050      1060      1070      1080
 D  A  S  R  D  A  R  N  A  E  E  K  A  K  K  A  I  T  D  A  A  M  M  A  E  E  L  K  K  E
GACGCCAGCCGGGATGCCAGGAATGCCGAGGAGAAAGCCAAGAAGGCCATCACAGACGCGGCCATGATGGCGGAGGAGCTGAAGAAGGAG
      1090      1100      1110      1120      1130      1140      1150      1160      1170
 Q  D  T  S  A  H  L  E  R  M  K  K  N  M  E  Q  T  V  K  D  L  Q  H  R  L  D  E  A  E  Q
CAGGACACCAGCGCCCACCTGGAGCGCATGAAGAAGAACATGGAGCAGACGGTCAAGGACCTGCAGCACCGCCTGGACGAGGCCGAGCAG
      1180      1190      1200      1210      1220      1230      1240      1250      1260
 L  A  L  K  G  G  K  K  Q  I  Q  K  L  E  T  R  I  R  E  L  E  F  E  L  E  G  E  Q  K  K
CTGGCGCTCAAGGGTGGCAAGAAGCAGATCCAGAAGCTGGAGACGCGGATCCGAGAGCTGGAGTTTGAGCTGGAAGGGGAGCAGAAGAAG
      1270      1280      1290      1300      1310      1320      1330      1340      1350
 N  A  E  S  V  K  G  L  R  K  Y  E  R  R  V  K  E  L  T  Y  Q  S  E  E  D  R  K  N  V  L
AACGCAGAATCTGTGAAGGGCTTGAGGAAGTATGAGCGGCGGGTCAAGGAGCTAACGTACCAGAGTGAAGAGGACAGGAAGAACGTGCTC
      1360      1370      1380      1390      1400      1410      1420      1430      1440
 R  L  Q  D  L  V  D  K  L  Q  V  K  V  K  S  Y  K  R  Q  A  E  E  A  D  E  Q  A  N  V  H
CGGCTGCAGGACCTGGTGGACAAGCTGCAAGTCAAGGTCAAGTCCTACAAGAGGCAGGCCGAGGAGGCGGATGAGCAAGCCAACGTGCAT
      1450      1460      1470      1480      1490      1500      1510      1520      1530
 L  S  K  F  R  K  A  Q  H  E  L  E  E  A  E  E  R  A  D  I  A  E  S  Q  V  N  K  L  R  A
CTCAGCAAGTTCCGGAAAGCGCAGCACGAGCTGGAGGAGGCGGAGGAGCGGGCCGACATCGCCGAGTCGCAAGTCAACAAGCTTCGGGCT
      1540      1550      1560      1570      1580      1590      1600      1610      1620
 K  T  R  D  F  T  S  S  R  M  V  V  H  E  S  E  E  *
AAAACCCGGGACTTCACCTCCAGCAGGATGGTTGTCCATGAGAGTGAAGAGTGAGCGAGCGTGTCCTGGGGCAGGACAGAGGACACGCAG
      1630      1640      1650      1660      1670      1680      1690      1700      1710

AATGCGTGTTTTCATGCTCCCTGCCCTTACACCTGACGTCCATGGGACCGAATTC
      1720      1730      1740      1750      1760
```

Fig. 1. Nucleotide and derived amino acid sequence of clone MHC 6–6. The termination codon is represented by (*).

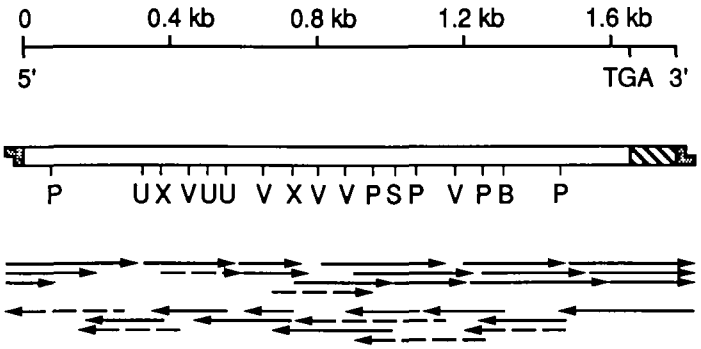

Fig. 2. Sequencing strategy used to sequence the LMM clone MHC 6–6 and restriction sites used in cloning manipulations involving MHC 6–6 and the expression vector pIN II A2. Filled arrows show regions sequenced with M13 universal primer using restriction fragments, whereas broken arrows indicate regions sequenced using primers based on preliminary sequence data. Restriction sites are identified by letters: B, *Bam*HI; P, *Pst*I; S, *Sac*I; U, *Sau*IIIa; V, *Pvu*II; X, *Xho*I. The *Eco*RI linkers are shaded and the 3'-untranslated region is crosshatched.

The open reading frame of this clone (designated MHC 6–6) encoded 557 amino acids extending to the C-terminal end of the myosin heavy chain. The derived amino acid sequence showed extensive homology with other skeletal muscle myosin heavy chain sequences (approximately 80%) and alignment of the sequence indicated that the N-terminal amino acid was 105 residues C-terminal to the start of LMM, as determined by Lu and Wong (1985). MHC 6–6, therefore, encoded over 80% of the LMM region.

Bacterial expression and purification of proteins

Restriction sites used in cloning manipulations involving MHC 6–6 and the expression vector pIN-II-A2 are shown in Fig. 2. To obtain in-frame translation of MHC 6–6, an *Eco*RI fragment from pUC18 encompassing the entire coding region and the 3' untranslated region of the MHC 6–6 cDNA, was subcloned into the *Eco*RI site of pIN-II-A2 (Fig. 3), to generate the vector pIN-II-MHC6–6, which expressed the entire open reading frame of MHC 6–6, with six amino acids of the lipoprotein fused to the N terminus. The polypeptide product of this vector was denoted LMMr

Other fragments were subcloned for expression using unique restriction sites within the myosin rod sequence. Blunt-end cutters facilitated the use of synthetic linker oligonucleotides to obtain translation in the correct phase or to insert termination codons. To delete part of the C terminus of LMMr, the pUC-MHC6–6 vector was cut at the unique *Rsa*I site at position 1408 of the MHC 6–6 sequence. Synthetic oligonucleotide linkers (New England Biolabs) with the sequence 5' CTAGTCTAGACTAG 3' were ligated to the ends of the fragment. This oligonucleotide contained the cleavage site for *Xba*I, and termination codons in all three phases. The fragment was then cut with *Xba*I, *Eco*RI and *Hin*dIII. The 1400 base pair *Eco*RI–*Rsa*I/*Xba*I fragment and the 200 base pair *Xba*I–*Hin*dIII fragment were ligated into *Eco*RI/*Hin*dIII-cut pIN-II-A2. The translational reading frame of this vector encodes the N-terminal 469 amino acids of LMMr, with four additional amino acids (Ser-Ser-Leu-Tyr) derived from the linker sequence on the 5' end. As these extra residues were uncharged and similar to those in this region of the original sequence it was unlikely that their presence would alter the solubility properties of the molecule substantially. This vector (pIN-II-MHC-N469) encoded a polypeptide lacking 92 amino acids from the C terminus of the myosin sequence. All mutants other than N469 had the same C terminus as the endogenous sequence.

An expression vector (pIN-IIMHC-C338) lacking the N-terminal 223 amino acids of MHC 6–6, but extending to its C terminus, was constructed as follows: pUC-MHC 6–6 was cut with *Fsp*I and then *Eco*RI linkers with the sequence 5' GGAATTCC 3'

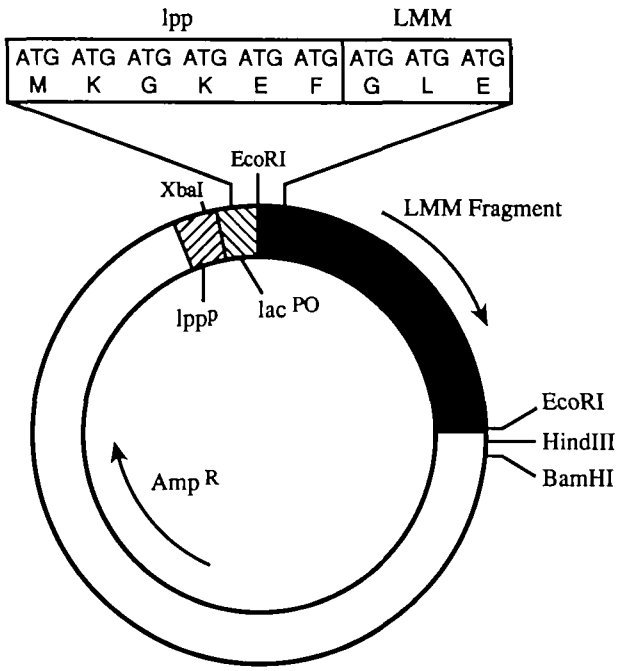

Fig. 3. Schematic representation of the lpp vector plasmid construct pIN-II-MHC 6.6 used to express the cDNA in *E. coli*. This construct produced a fusion protein with six residues from the *E. coli* lipoprotein located at the LMM N terminus. Transcription was from the lpp promotor and sequences from the *lac* operon rendered expression inducible by IPTG.

were ligated to the fragment. This fragment was then cut with *Eco*RI and the 1095 base pair fragment ligated into *Eco*RI-cut pIN-II-A2. The orientation of the insert was determined by cutting with *Hin*dIII, since only the correct orientation gave the expected 165 base pair fragment. The expressed protein had the same N-terminal amino acids derived from the lipoprotein sequence as LMMr, with Ala224 the first residue derived from the myosin sequence.

Trials for expression of new constructs were carried out on a 10 ml scale whereas preparative scale expression of recombinant proteins was carried out in 1 l bacterial cultures. When the cultures had reached an absorbance at 600 nm of 0.1–0.2, expression of the protein was induced by addition of isopropyl-β-D-thiogalactoside (IPTG) to a concentration of 125 $\mu g\,ml^{-1}$. The cultures were then grown for a further 2–3 h and the cells harvested by centrifugation at 3000 *g* for 10 min. The pellet was washed with 20 ml of cold 100 mM NaCl, 20 mM Tris–HCl (pH 8.0), 1 mM EDTA, then rapidly frozen in liquid nitrogen and kept at −20°C overnight. The pellet was then thawed, resuspended in 10 ml of ice-cold lysis buffer (50 mM Tris–HCl (pH 8.0), 1 mM EDTA, 25% (w/v) sucrose) and lysozyme added to $1\,g\,l^{-1}$. Lysis was allowed to proceed for 10 min on ice, then an equal volume of 2 M NaCl in lysis buffer was added and the cells were sonicated using a Heat Systems Ultrasonics sonicator with a 1 cm probe. We then took advantage of the fact that light meromyosin, like some other fibrous proteins such as tropomyosin, is stable against heat treatment (see, for example, Hitchcock-DeGregori and Heald, 1987). Immediately after sonication, the homogenate was heated to 95–100°C in a microwave oven for 5 min, cooled on ice for 30 min, and centrifuged at 120 000 *g* for 1 h at 4°C.

A substantial enrichment for LMMr was achieved by this step, so that it constituted more than 60% of the total protein in the sample. The expressed myosin fragments were fractionated by ammonium sulphate precipitation between 25 and 40% saturation to remove the bulk of the nucleic acids. The 40%

ammonium sulphate pellet was dispersed in 40 mM sodium pyrophosphate (pH 8.0), 1 mM EDTA and purified by anion exchange chromatography on a Mono-Q FPLC column (Pharmacia) in the same buffer. The myosin fragments were eluted with a linear gradient of 0 to 500 mM NaCl, the fragments eluting at about 300 mM NaCl. This purification protocol had the advantage that it made no assumptions about the solubility properties of the expressed protein, and so could be used for wild-type LMM and also for soluble mutants. Although initially some problems were encountered with partial proteolysis of some expressed proteins, this problem was reduced to negligible proportions when this protocol was followed.

Skip residue deletions

Skip residues were assigned on the basis of homology with other skeletal muscle myosin heavy chain sequences to Glu200 and Gly425. Oligonucleotide-directed site-specific mutagenesis was carried out by the method of Carter *et al.* (1985) as described by Quinlan *et al.* (1989). Restriction fragments were subcloned to M13mp18 or mp19 for mutagenesis, and mutants were checked by sequencing the whole insert before subcloning back into the expression vector. An oligonucleotide with the sequence 5′ GCCGAGAAGGACGAGATGGAGCAGCTGAAG 3′ was used to delete Glu200, to generate LMM-NSKD. An oligonucleotide with the sequence 5′ CAGCTGGCGCTCAAGGGCAAGAAGCAGATC 3′ was used to delete Gly425, to generate LMM-CSKD.

Sedimentation techniques

A sedimentation assay was used to determine the concentrations of monomer and polymer for the mutant and wild-type proteins. Protein at variable concentrations in high salt buffer (0.5 M NaCl, 5 mM sodium phosphate (pH 7.0)) was diluted 10-fold into lower salt buffers, and left for 90 min at room temperature. The mixture was centrifuged for 10 min at 10 000 *g*. The absorbance at 280 nm was used to measure protein concentration. Generally a final total protein concentration of 0.1 mg ml^{-1} was employed, although for the very soluble N469 mutant, concentrations as high as 4 mg ml^{-1} were used.

Paracrystals were found in the precipitates formed in most low ionic strength buffers, up to 150 mM NaCl, within the pH range 7–8. Below pH 7 the material tended to precipitate as large aggregates that showed little sign of striations in negative stain. Above pH 8 the precipitate tended to consist of much smaller aggregates that appeared less well ordered. Paracrystals with the clearest banding pattern were obtained by dialysis against 50–120 mM NaCl, at a pH of 7.0–7.5, with sodium phosphate (5 mM) or bis-Tris propane (10 mM) as buffer. The latter seemed to give more reproducible results. Low concentrations of Mg^{2+} or Ca^{2+} had little effect on the appearance of the paracrystals. Paracrystals prepared by dilution into the appropriate buffer appeared substantially less well ordered than those formed by dialysis.

Electron microscopy

Specimens were negatively stained with uranyl acetate. Material was adsorbed onto carbon-coated, 400-mesh grids by applying a 5 μl drop of sample for 10–20 s. The grid was rinsed with 2–3 drops of buffer, before staining with 6–8 drops of 2% aqueous uranyl acetate. For more consistent staining, the Valentine *et al.* (1968) method was used. For shadowing, samples were sprayed, in an appropriate buffer containing 50% glycerol, onto freshly cleaved mica, and then shadowing with platinum–carbon in a Cressington (Watford, Herts.) CFE50 freeze etching device at a nominal angle of 10 degrees, to a thickness of 1–1.5 nm.

Specimens were examined in Philips EM400 and CM 12 electron microscopes operated at 80 or 100 kV, usually with an anti-contamination device. Some images of negatively stained specimens were recorded under low-dose conditions. Images were recorded on Kodak SO163 or Agfa Scientia 23D56 sheet film and developed using Ilford PQ universal developer. Microscope magnifications were calibrated by reference to the 2.81 nm spacing of negatively stained sheaths of *Methanospirillum hugatei* (Stewart *et al.* 1985*a*).

Results

Expression and characterization of rabbit LMM and deletion mutants

We expressed in *E. coli* a fragment of rabbit skeletal muscle myosin cDNA (MHC 6–6) corresponding broadly to the LMM domain of the rod. The fragment was 1765 bp long and included a coding region of 1664 bp. The 5′ nucleotide of the fragment corresponded to nucleotide 4263 of the chick sequence of Molina *et al.* (1987), and nucleotide 377 of the rabbit fast skeletal muscle sequence (Maeda *et al.* 1987). The derived amino acid sequence of MHC 6–6 is shown in Fig. 1. The clone encoded 557 amino acids, giving a predicted molecular weight of 64.59×10^3. The alignment of MHC 6–6 with other myosin sequences showed that its N terminus lay 105 amino acids C-terminal to the S2-LMM junction as determined by Lu and Wong (1985), so that MHC 6–6 encoded most of the LMM region of myosin.

The protein encoded by the MHC 6–6 cDNA, denoted LMMr, was expressed to high levels using the pIN vector system. It was the most prominent band on gels of whole cell lysates and probably constituted between 5 and 10% of total cell protein. On SDS–PAGE LMMr migrated with an apparent molecular weight of 70×10^3 M_r (Fig. 4, below) and gave a positive reaction with a broad-spectrum myosin antibody on Western blots. The sequence of LMMr predicted a protein product with a molecular weight of 64.6×10^3 M_r. The slightly higher M_r of 70×10^3 determined by SDS–PAGE was not surprising, as it is commonly found that fibrous proteins have an anomalously slow mobility on gels (see, for example, Sutoh *et al.* 1978; Kaufman *et al.* 1984; Noegel *et al.* 1989). The N-terminal sequence of the 70×10^3 M_r band cut from a SDS–PAGE gel was

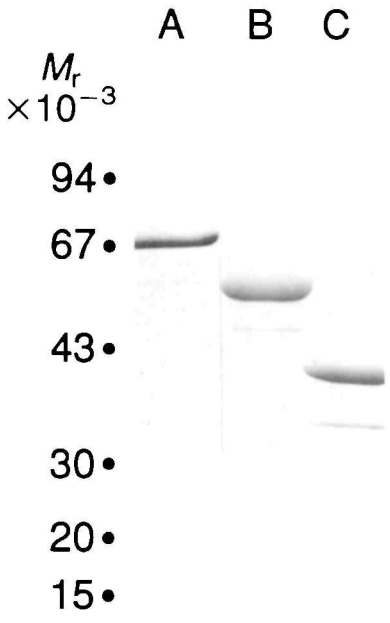

Fig. 4. SDS–PAGE of expressed myosin rod fragments. (A) LMMr; (B) LMM-N469; (C) LMM-C388. The positions of molecular weight markers (phosphorylase *b*, 94×10^3 M_r; albumin, 67×10^3 M_r; ovalbumin, 43×10^3 M_r; carbonic anhydrase, 30×10^3 M_r; trypsin inhibitor, 20×10^3 M_r; alpha-lactalbumin, 14.5×10^3 M_r) are shown on the left.

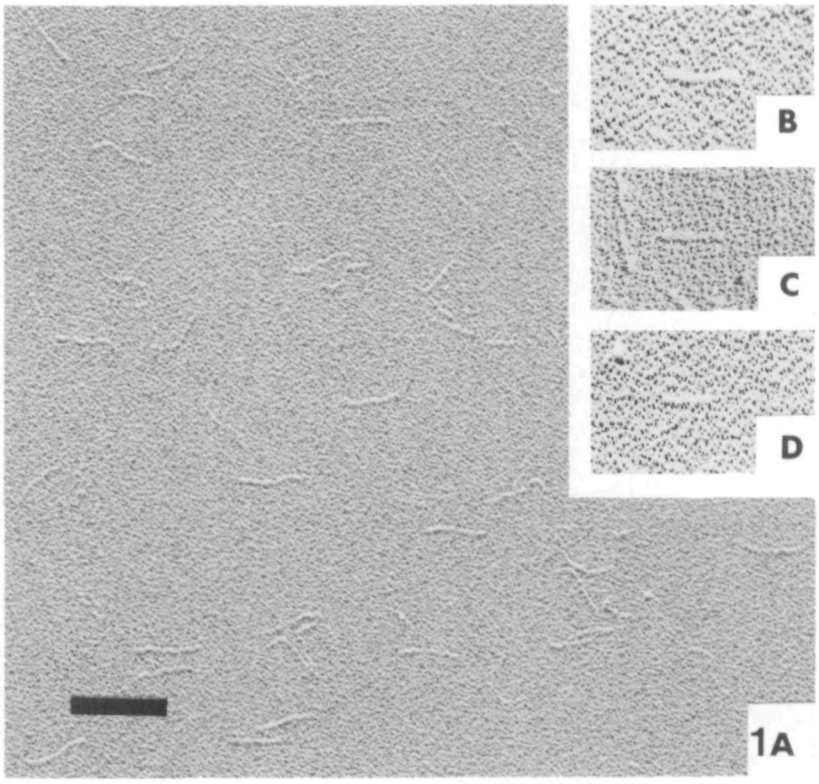

Fig. 5. (A) Electron micrograph of field of shadowed LMMr molecules. Inset: gallery of electron micrographs of shadowed molecules of the different constructs: (B) LMMr; (C) LMM-N469; (D) LMM-C338. Bar, 100 nm.

MKGKEFGLEEAKKKLAQ, identical to the sequence predicted for the correct expression construct, with six amino acids of the lipoprotein sequence preceding the LMMr sequence (see Fig. 3). Disulphide crosslinking (Stewart, 1975) doubled the M_r on non-reducing gels confirming the formation of a coiled-coil in which the two chains were in register as seen with proteolytic fragments of myosin rod (Stewart, 1982). The final yield of pure LMM 6–6 was usually 1–2 mg per litre of bacterial culture.

Molecules were constructed with deletions from either the N or C terminus of LMMr and the resultant proteins were expressed and purified in the same way as the full-length material. These constructs expressed protein to levels comparable with LMMr Fig. 4 shows samples of purified material analysed by SDS–PAGE. The expressed fragments of LMM-C338 and LMM-N469 ran with apparent molecular weights of 39 and 60×10^3 M_r, respectively. A small quantity of contaminating material was present in some preparations. Analysis of the sequence predicted molecular weights of 39×10^3 M_r for LMM-C338 and 54×10^3 M_r for LMM-N469. Mutants with the skip residues deleted were also expressed to similar levels, and migrated on gels in the same position as LMMr

Electron microscopy demonstrated that these proteins had a distinctive rod-like morphology (Fig. 5), consistent with an extended coiled-coil conformation (see, for example, Stewart and Edwards, 1984). A montage of shadowed molecules of LMMr, LMM-C338 and LMM-N469 is shown in Fig. 5 (inset). This clearly demonstrated the gradation in size between the different molecules. The molecules all had the same rod-like appearance. When appropriate concentrations were chosen (usually 50 $\mu g\,ml^{-1}$), individual molecules well separated from their neighbours were commonly observed (Fig. 5), which enabled the lengths of the particles to be measured. The lengths of a sample of molecules of each fragment were measured and the data are shown in Table 1. All the fragments showed good agreement between the predicted and measured lengths of the molecules.

Solubility properties of expressed proteins

LMMr was largely insoluble at low ionic strength (50 mM NaCl) at pH 7.0. As the pH and ionic strength were increased, the solubility also increased (Fig. 6) in the same general way that was observed for proteolytic rabbit skeletal muscle LMM (Katsura and Noda, 1971, 1973). The critical concentration for assembly of LMMr, as assessed by the concentration of protein remaining in the supernatant after centrifugation, was approximately 25 $\mu g\,ml^{-1}$ at pH 7.0, and 50 mM ionic strength, rising to about 65 $\mu g\,ml^{-1}$ at pH 8.0. LMM-C338 was, if anything, slightly

Table 1. *Length data (predicted and observed) for LMM fragments*

	$M_r\times10^{-3}$		Length (nm)			
				Observed		
Fragment	Predicted	Observed	Predicted	Mean	S.D.	n
LMMr	64.6	70	80	79	4	81
LMM-N469	51	60	69	67	4	20
LMM-C338	39	39	50	51	4	20
LMM-CSKD	64.6	70	80	79	4	20

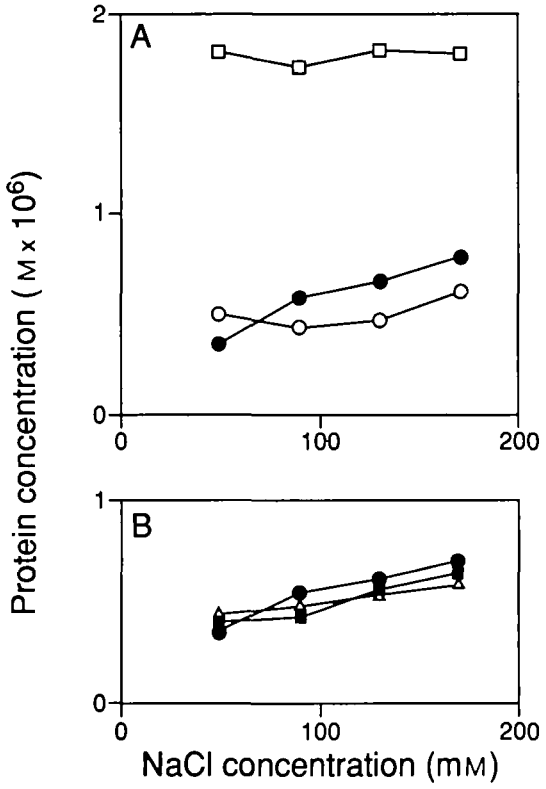

Fig. 6. Solubility of the different LMM constructs as a function of ionic strength. (A) Deletion mutant constructs: (●) LMMr; (□) LMM-N469; (○) LMM-C338. Although the solubility behaviour of LMM-C338 was almost indistinguishable from that of the wild-type LMMr, the solubility of the construct with 92 residues deleted from the C terminus, LMM-N469, was strikingly higher. In this experiment, the total protein concentration employed for the assay was 0.1 mg ml^{-1} and all samples except LMM-N469 precipitated to some extent under these conditions. However, LMM-N469 was completely soluble even when concentrations of 4 mg ml^{-1} (80 μM) were employed. (B) Skip residue deletion constructs: (●) LMMr; (△) CSKD; (■) NSKD. Deletion of either skip residue had a negligible effect on the solubility properties of the construct.

less soluble than LMM 6–6 over the range of buffer conditions tested (Fig. 6A). Both LMMr and LMM-C338 were fully soluble at higher ionic strength (above 0.3 M). However, LMM-N469 was completely different in its solubility properties at low ionic strength. The concentration of LMM-N469 measured in the supernatant after dialysis was indistinguishable from the starting total protein concentration over the range of concentrations tested. This fragment remained soluble up to concentrations of 4 mg ml^{-1}, which was two orders of magnitude higher than the solubility observed with LMMr The constructs from which skip residues had been deleted (LMM-NSKD and LMM-CSKD) were indistinguishable from LMMr in their solubility properties (Fig. 6B).

Formation of ordered aggregates

The material precipitated under low ionic strength conditions was examined by electron microscopy using negative staining. LMMr formed large aggregates that were sometimes up to many micrometres long. They generally had a spindle-like shape (best seen at lower magnifications, e.g. Fig. 7), and often tended to be clumped together. Larger specimens tended to be too thick to obtain satisfactory electron micrographs. They showed a clear pattern of light and dark stripes perpendicular to the long axis of the paracrystal (Fig. 7). In addition, a number of crystalline sheets were also produced under these conditions.

LMMr paracrystals. There was a substantial degree of polymorphism observed in the paracrystalline aggregates formed when LMMr was dialysed to low (near physiological) ionic strength. The most common pattern seen on well-isolated paracrystals showed a simple pattern of narrow dark stripes separated by wide lighter bands (Fig. 7) with an axial repeat of 43 nm. The dark stripes were about 6 nm wide, and the light bands were about 36–37 nm wide. Most paracrystals appeared to have dyad symmetry, which indicated an antiparallel packing of LMM 6–6 molecules. Occasionally polar paracrystals, also with a 43-nm axial repeat, were observed (Fig. 8), but these were much less common than the other forms. One could often see that the ends of most types of paracrystal, and places where the width of the paracrystals decreased in a series of steps, coincided with the end of the light band (Figs 7 and 8).

These results indicated that the packing of LMM 6–6 molecules in the paracrystals could best be explained by the type of gap–overlap structure analogous to that found with paracrystals of paramyosin (Cohen *et al.* 1971), and of the intermediate filament protein GFAP (Stewart *et al.* 1989), as well as with LMM (see, for example, Katsura and Noda, 1973; Bennett, 1981). The type of structure proposed for the basic 43-nm LMM 6–6 paracrystals is illustrated in Fig. 9. Molecules overlap their nearest neighbours by 43 nm. The 6–7 nm dark band means that the molecules must be 6–7 nm shorter than twice the molecular overlap. This predicts a molecular length of 79–80 nm, which is in excellent agreement with the length predicted from the sequence (80 nm) and that measured from shadowed preparations (79 nm, Table 1). The molecular orientation (that is to say, whether the longer overlap was between N or C termini in dyad paracrystals) could not be determined by inspection of the paracrystals.

Areas displaying a more complex pattern (for example, see Fig. 10) were quite common. Often these occurred where paracrystals with the simple 43 nm pattern appeared to overlap. Staggers of one third of the fundamental repeat (14.3 nm) were most frequent (Fig. 10), in which dark stripes were separated by light bands alternately 7 and 21 nm long. In addition, a stagger that appeared to be closer to half the repeat (approximately 20 nm) was observed more rarely (for example, part of the paracrystal shown in Fig. 10). Many of the polymorphic paracrystal forms that we observed appeared to be generated by superposition of axially staggered units of the fundamental repeat. The paracrystal form with the two dark lines separated by one third of the fundamental repeat corresponded to interactions between molecules staggered by 14.3 nm, which is the axial separation of myosin molecules in muscle thick filaments. These polymorphic forms suggest that in addition to the 43 nm spacing seen in the simple paracrystal form, interactions between molecules based on a 14.3 nm repeat are also possible.

Two-dimensional crystalline sheets. In addition to the paracrystals described above, we sometimes found thin crystalline sheets in preparations of LMMr dialysed to low ionic strength. The most common pattern observed (Fig. 11) was one with an axial repeat of 43 nm in which two dark lines were separated by 19 nm (just under half the repeat distance). At high magnification (Fig. 12) it

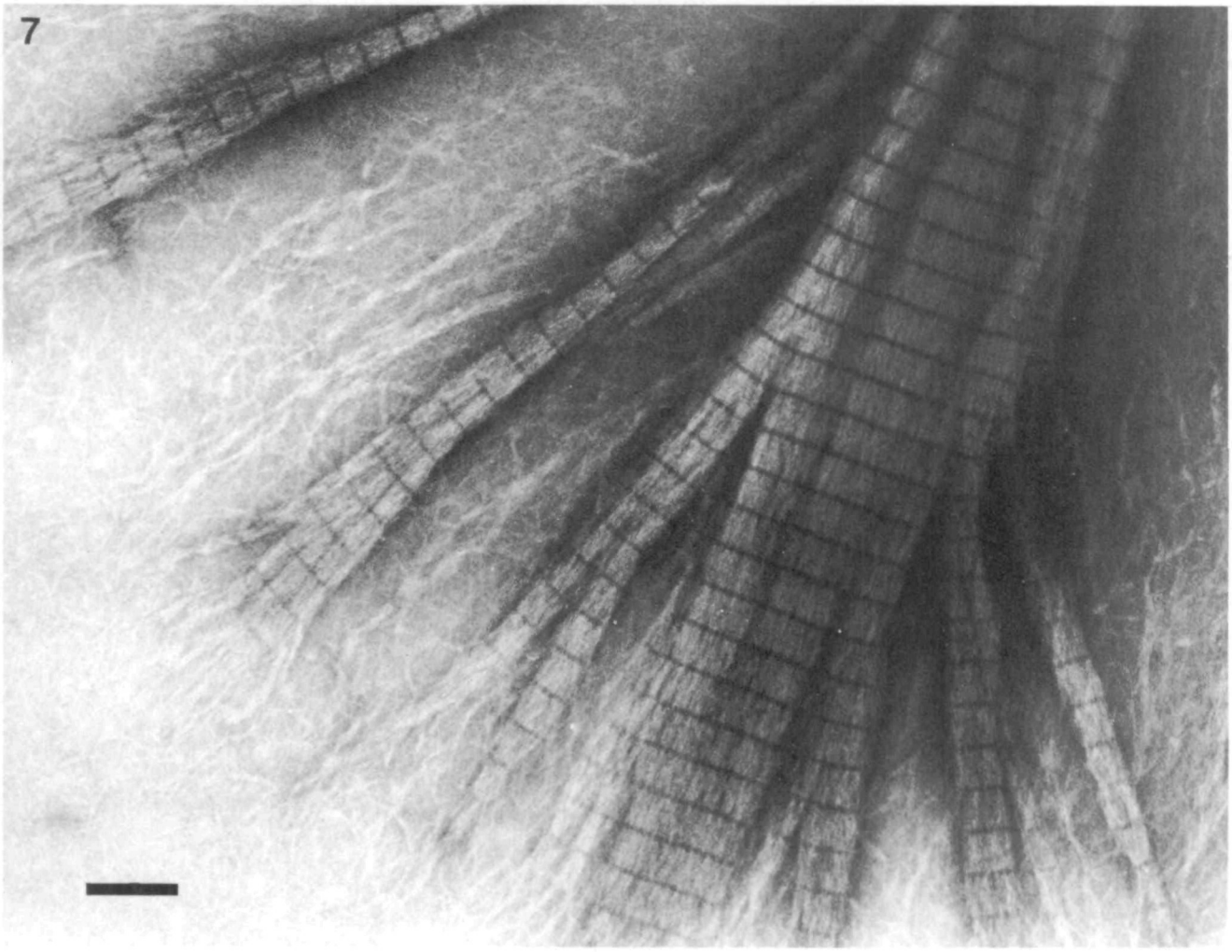

Fig. 7. Electron micrograph of LMMr paracrystals negatively stained with uranyl acetate showing the pattern most commonly observed. These paracrystals have a single prominent dark line spaced axially every 43 nm and appear to have dyad symmetry. Note that at the fraying ends of the paracrystals, the dark line appears to coincide with steps in the structure and paracrystal ends. This observation indicates that the molecular ends are located at this line. Bar, 100 nm.

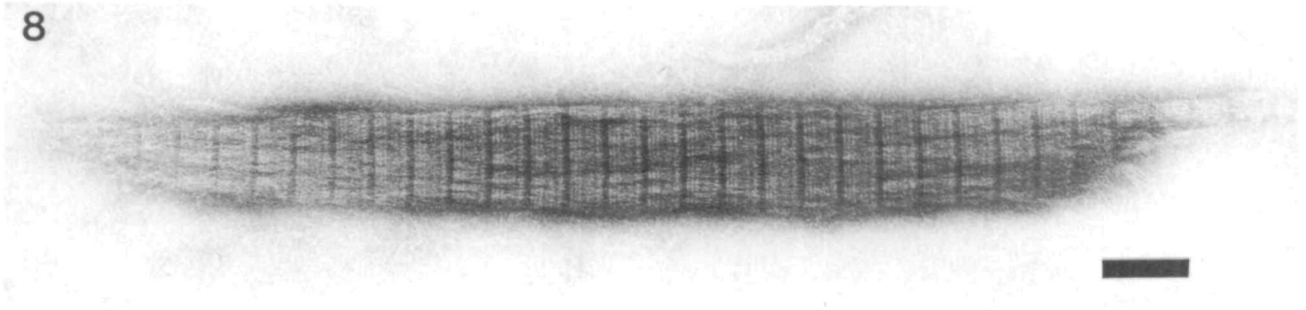

Fig. 8. Polar paracrystal of LMMr. This pattern clearly does not have dyad symmetry, and so indicates that the molecules within the paracrystal are arranged parallel to one another. This polar pattern was observed much less frequently than the dyad pattern shown in Fig. 7. Bar, 100 nm.

could be seen that each of these dark stripes perpendicular to the long axis of the crystal was made up of alternating dark and light lines separated by about 4 nm. Optical diffraction (Fig. 12, inset) showed characteristic row lines at a lateral spacing of 3.6 nm sampled by the 43-nm axial repeat, confirming that the molecules were arranged on a two-dimensional crystalline lattice.

Aggregates with LMM mutants. Although LMM-C338 was precipitated under conditions of low ionic strength, no paracrystalline structures were observed under the con-

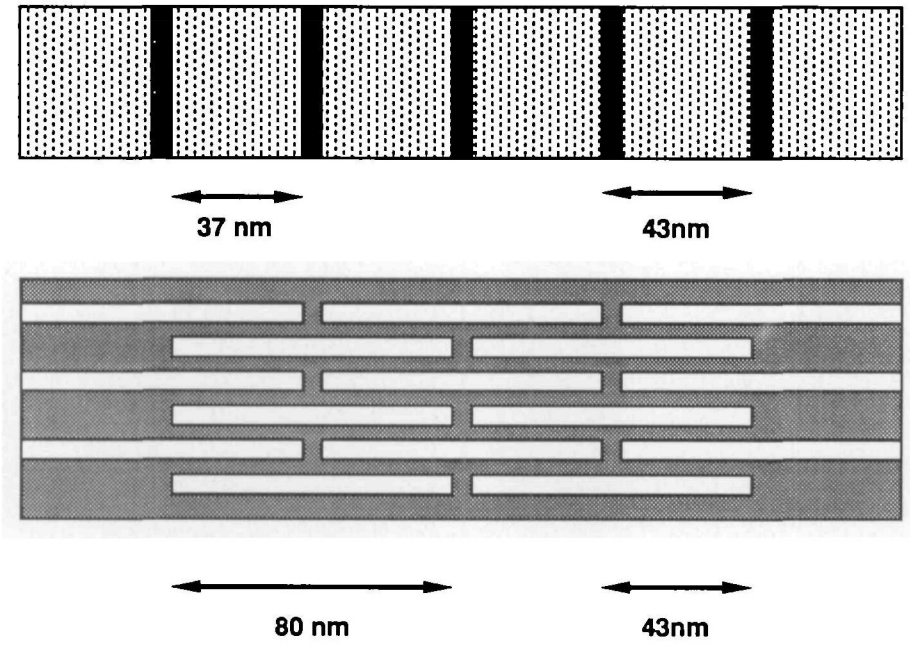

Fig. 9. Schematic representation of the molecular arrangement in the LMMr paracrystals with a 43 nm axial repeat. The individual 80 nm long molecules are staggered by 43 nm, but, because the molecular length is less than twice the axial repeat, there is a roughly 6 nm gap between successive molecules. Because stain will penetrate this gap, it appears in darker contrast to the surrounding material. Consequently, the negatively stained pattern is dominated by a dark line every 43 nm.

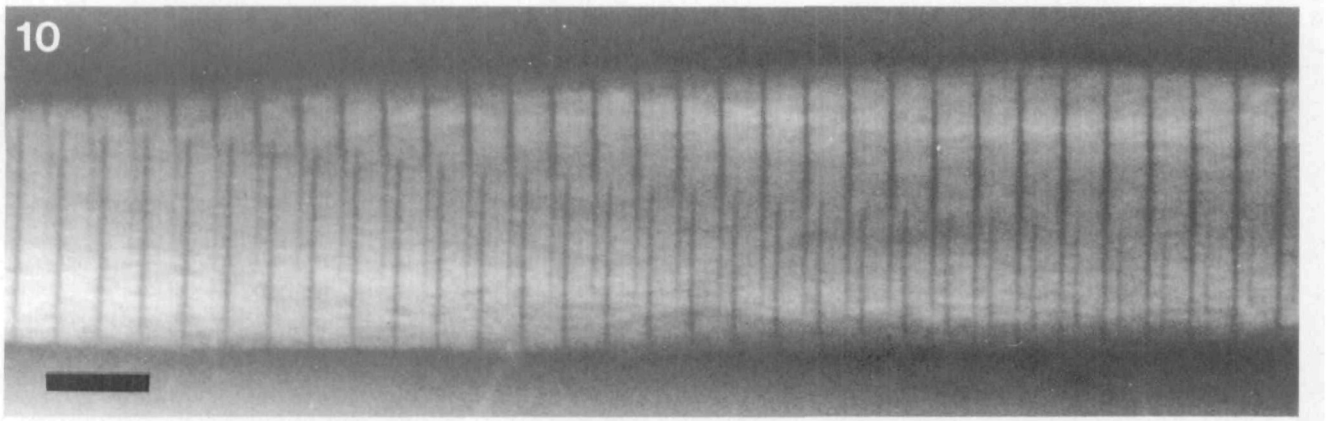

Fig. 10. Electron micrograph of a negatively stained LMMr paracrystal showing an example of the more complex patterns formed by superposition of two areas having the simpler pattern that is illustrated in Fig. 7. Bar, 100 nm.

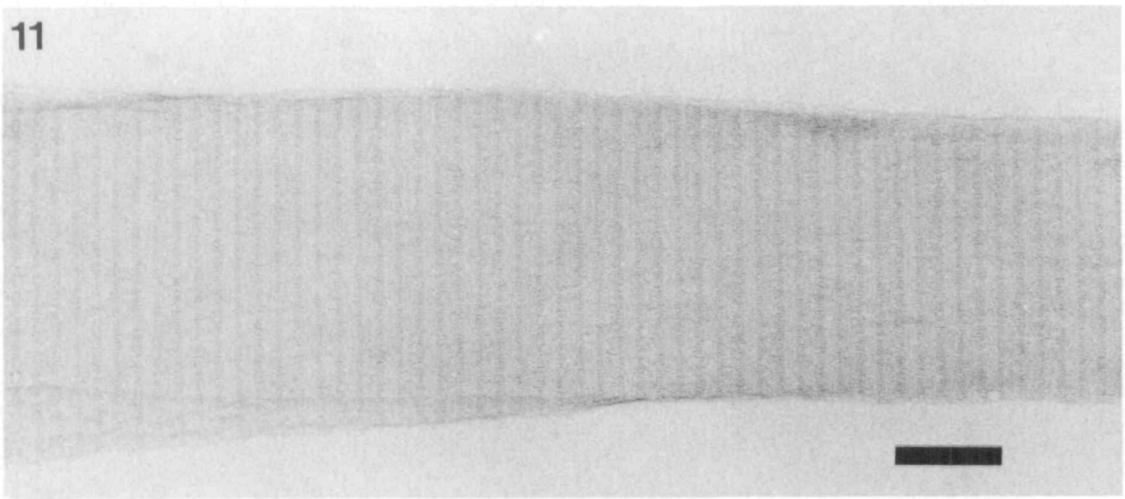

Fig. 11. Electron micrograph of a negatively stained crystalline sheet of LMMr, showing an axial banding pattern dominated by two dark stripes, separated by 19 nm, every 43 nm. Bar, 100 nm.

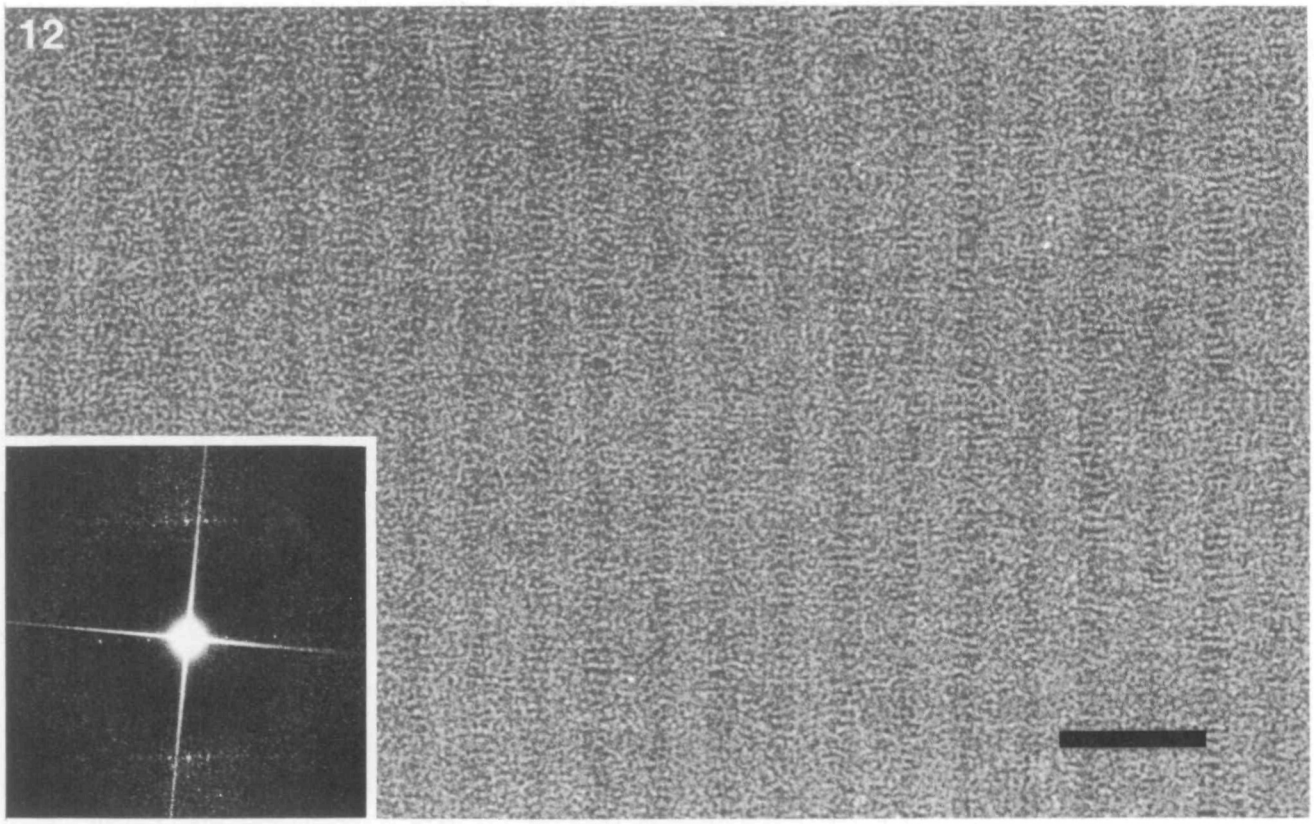

Fig. 12. Higher-magnification electron micrograph of part of the negatively stained crystal illustrated in Fig. 11. When viewed from the side, it is clear that each dark stripe is composed of alternate light and dark lines separated laterally by about 4 nm. Bar, 50 nm. Inset: optical diffraction pattern showing prominent off-axial row lines corresponding to a lateral spacing of 3.6 nm. These row lines, and the meridian, are sampled by the finer repeat due to the 43 nm axial banding pattern.

ditions examined. The material formed short narrow filaments that tended to aggregate into larger tangles showing no sign of any degree of order (Fig. 13). LMM-N469 was, of course, soluble under these conditions and so did not form paracrystals. Both LMM-NSKD and LMM-CSKD formed paracrystals. Those formed by LMM-CSKD were very similar to those formed by wild-type material. The aggregates had the same spindle-like appearance, and a strong 43 nm axial banding pattern was observed (Fig. 14A). Both the polymorphic paracrystal forms observed in wild-type material (having staggers of one third and close to half the 43 nm repeat) were observed in aggregates of LMM-CSKD. LMM-NSKD formed aggregates with the same general morphology as LMM 6–6 and LMM-CSK (Fig. 14B). The aggregates had a strong axial banding pattern. Generally, the banding pattern was the same as that observed in a polymorphic forms of the LMMr paracrystals, with the 43 repeats staggered by 20 nm, close to half the repeat (Fig. 14). Areas displaying the simple 43 nm repeat were rare and small, and the other polymorphic form (staggered by one third of the 43 nm repeat) was not observed.

Discussion

Bacterial expression of fragments of myosin rod provides a powerful means by which to study the molecular interactions that are involved in myosin assembly. Defined fragments can be expressed in *E. coli*, enabling the regions of the rod that are important for assembly to be defined (Sinard *et al.* 1990; O'Halloran *et al.* 1990). Such fragments may have potential value in crystallization studies in order to obtain high-resolution information about the myosin rod. The technique of site-directed mutagenesis enables precise changes to be made in the rod sequence, and the effect of these changes on the *in vitro* assembly properties of the rod fragments can then be examined.

We have produced substantial quantities of myosin rod fragments by expression of a myosin heavy chain cDNA in *E. coli* using the pIN vector system (Nakamura and Inouye, 1982), which has also been used by Leinwand *et al.* (1989) to express myosin fragments. Expression levels were generally high, and the protein could be purified to homogeneity using a protocol based on the heat stability of LMM. The length of the molecules was measured in electron micrographs of shadowed preparations and found to correlate well with the length predicted from an alpha-helical coiled-coil structure. LMMr and LMM-C338 had solubility properties similar to those of proteolytic LMM and the critical concentration for their assembly varied in a similar manner with pH and ionic strength. Moreover, in physiological ionic strength conditions, the recombinant material formed paracrystals with the 43 nm and 14.3 nm axial repeats observed in aggregates obtained from myosin rod and a range of its proteolytic fragments (Kendrick-Jones *et al.* 1971; Katsura and Noda, 1973; Safer and Pepe, 1978; Yagi and Offer, 1981; Quinlan and Stewart, 1987; Ward and Bennett, 1989). These data support strongly the conclusion that we have produced material with near native propeties. Maeda *et al.* (1989)

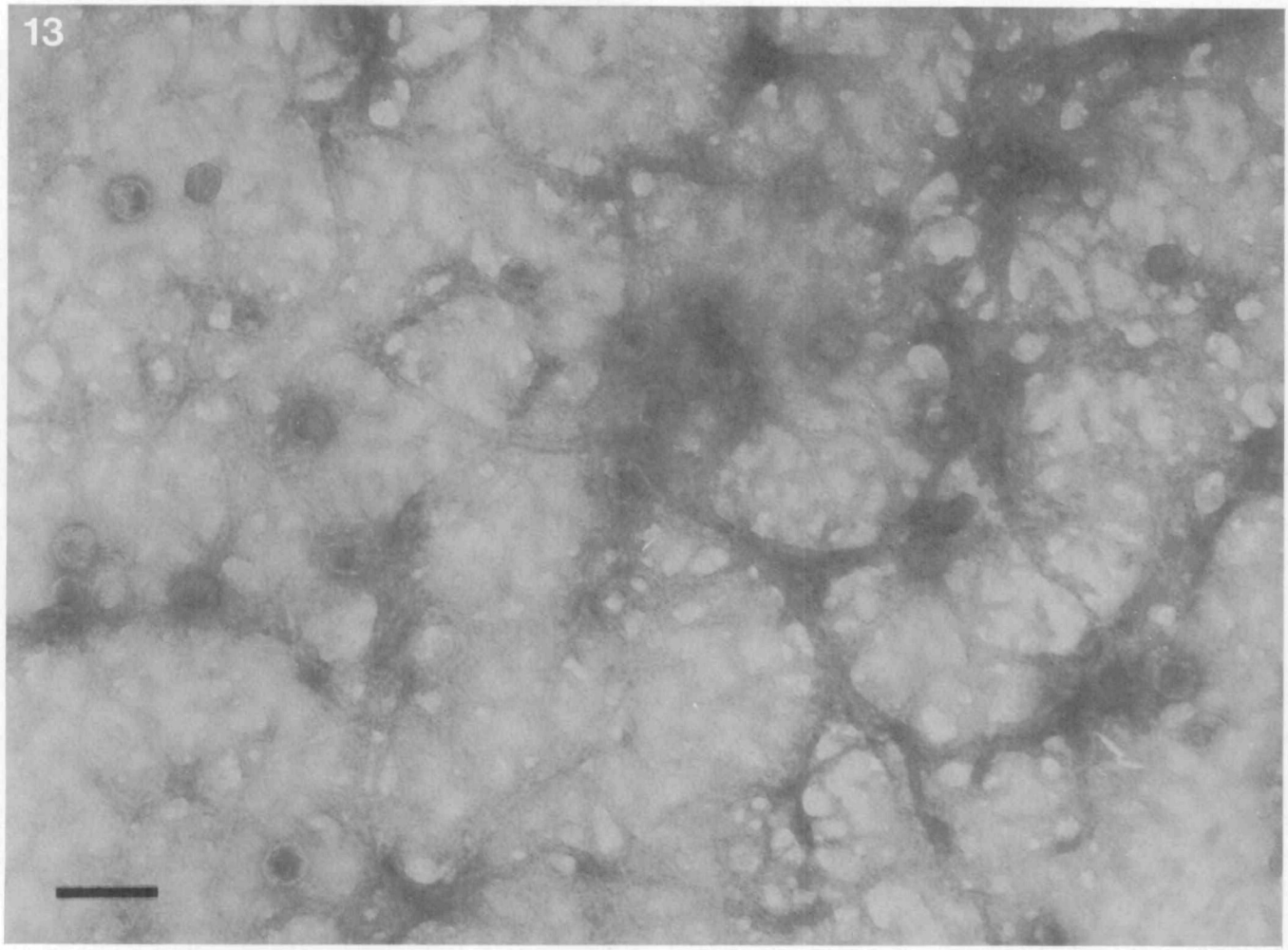

Fig. 13. Electron micrograph showing a negatively stained field of the disordered fibrous aggregates formed when LMM-C338 precipitated at low ionic strength. Bar, 100 nm.

have also shown the formation of paracrystals from expressed rabbit myosin LMM.

The type of gap–overlap structural model we have proposed for the LMMr paracrystals is very similar to models proposed for paracrystals prepared from proteolytic LMM, such as those described by Katsura and Noda (1973) and Bennett (1981). Although in these studies using proteolytic LMM, the precise staining pattern depended on the extent of proteolysis, all the patterns observed were based on molecular staggers of 43 nm or 14.3, similar to those we observed with LMMr Moreover, in many instances it was clear that the 14.3 nm repeat was derived from an overlapping of 43 nm repeats in the same way as is illustrated in Fig. 10. Multiple polymorphic forms have been frequently observed in low ionic strength precipitates of proteolytic LMM (see, for example, Katsura and Noda, 1973, for a particularly well-documented study). The observation of such polymorphism in the more homogeneous LMMr prepartions indicated that the polymorphism derived from an intrinsic property of the myosin rod sequence, rather than the presence of a range of proteolytic fragments. Such a polymorphism would be consistent with the range of different molecular interaction geometries that myosin rods probably have to adopt when forming muscle thick filaments (see Atkinson and Stewart, 1991). The polar paracrystals (Fig. 8) and two-dimensional sheets (Fig. 11) represent structures only rarely seen with myosin rod fragments (in which molecules are usually arranged antiparallel). It may be that the formation of these rare aggregates was a consequence of the high degree of homogeniety of the preparations obtained by expression in *E. coli*. Previous examples of sheets, for example, have been obtained by using cyanogen bromide cleavage (Young *et al.* 1972) or very carefully controlled proteolysis to yield very homogeneous material (Quinlan and Stewart, 1987).

In addition to the paracrystalline arrays commonly seen with proteolytic myosin rod fragments, we were also able to obtain two-dimensionally ordered crystalline sheets with the full-length recombinant LMM. In addition to the characteristic 43 nm axial repeat, these crystals had a lateral spacing of 3.6 nm or slightly less than twice the expected diameter of an alpha-helical coiled-coil (Fraser and MacRae, 1973). This lateral spacing indicated that the fundamental packing unit in these crystals was probably a molecular dimer, which may be important because a range of data indicates that myosin dimers are a key intermediate in thick filament assembly (Davis, 1988; Davis *et al.* 1982; Godfrey and Harrington, 1970*b*). These crystals may enable the molecular interaction geometry both with the dimer and between dimers to be established using computer-based image processing (see, Stewart, 1988) as

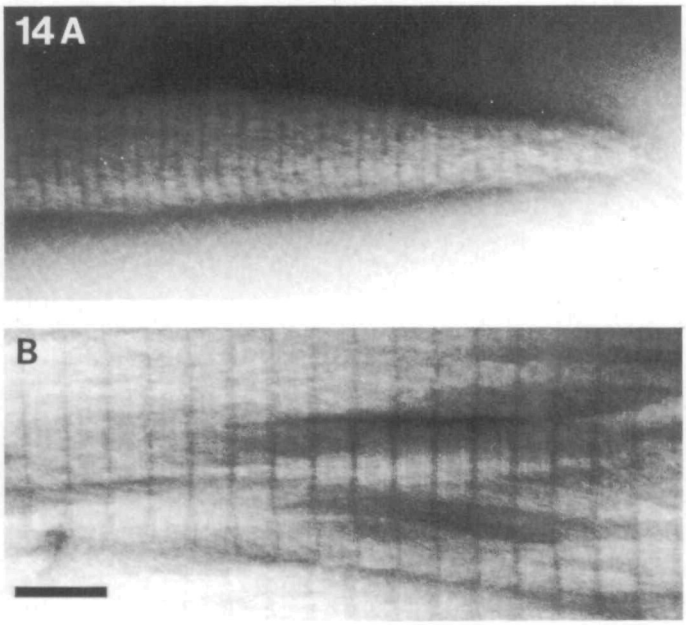

Fig. 14. Electron micrographs of negatively stained paracrystals formed from constructs from which skip residues had been deleted. (A) CSKD parac rystals; (B) NSKD paracrystals. Bar, 100 nm.

has been employed, for example, on crystalline tubes of the rod long S-2 fragment (Quinlan and Stewart, 1987). Crystal formation is comparatively rare with fibrous proteins and it is likely that our success with the expressed material was due to its high homogeniety, resulting not only from its not having been prepared by proteolysis, but also from its being only one isotype compared with the several that are usually prepared in myosin isolated from whole muscle.

Deletion of either skip residue in the LMMr molecule had no apparent effect on the solubility of the molecules, or on the structure of the paracrystals formed. This implied that the skip residues do not play a significant part in the molecular interactions that are responsible for the precipitation of LMM at low ionic strength. Apparently the structure of the coiled-coil was not perturbed by the presence of the skip residues, at least to the extent necessary to be detected by producing any change in the staining pattern or morphology of the aggregates formed. This observation would argue against the sort of distortion proposed by McLachlan and Karn (1982). Offer (1990) has argued that the positions of the skip residues are correlated with more flexible regions in the myosin rod. If this is the case, it may be that their role is to allow the rod to be distorted slightly to permit myosin molecules to pack correctly in the (non-crystalline) thick filament. Such distortion would probably not be important in the formation of paracrystalline and crystalline assemblies that characterize the aggregates formed by LMMr

The solubility properties of the different expressed fragments have important implications for the interactions between molecules that are important for assembly. In contrast to the very similar solubility properties of LMMr and LMM-C338, the deletion of the C-terminal region in LMM-N469 radically altered the solubility properties of this mutant. This mutant remained soluble over a range of pH and ionic strength conditions. The difference in solubility between LMM-N469 and LMM-C338 cannot be accounted for merely by the difference in their length, since although LMM-C338 is a shorter molecule than LMM-N469, it is less soluble. If the strength of the interaction was proportional to the length of the molecule (as might be anticipated if the major interaction between molecules derived from complementation of the zones of alternating positive and negative charge), then one would expect solubility to reduce with molecular length. Because the converse was observed, our results instead indicate that a specific part of the molecule, the C terminus, exerts a disproportionate influence on solubility and molecular interactions. The large increase in solubility that we observed by deletion of a small region at the C terminus of LMMr confirmed the findings obtained using proteolytic fragments of both skeletal muscle and smooth muscle myosin (Nyitray *et al.* 1983; Cross and Vandekerckhove, 1986) that indicated that an important determinant of LMM solubility lay near the C terminus of the molecule. A similar effect of deleting the C terminus of the molecule has been found using fusion proteins derived from *Acanthamoeba* myosin II (Sinard *et al.* 1990). In contrast, the 'solubility domain' of *Dictyostelium* myosin lies much further from the C terminus of the molecule (O'Halloran *et al.* 1990). Nyitray *et al.* (1983) found that a proteolytic fragment of LMM lacking as many as 129 amino acids remained insoluble at low ionic strength, and that it was necessary to remove a further 5×10^3 M_r before fragments became soluble. The discrepancy between this result and our own, which indicates that only 92 amino acids need be removed to alter the solubility drastically, may be due to the inhomogeneity of proteolytic preparations, or perhaps to differences between myosin isoforms.

A region near the C terminus of the myosin molecule, then, clearly has a vital role in its assembly. The molecular basis for this, however, remains obscure. The sequence appears not to be significantly different from that of the rest of the myosin rod, although a concentration of basic residues does mean that overall there is a dipole in the charged residues. Dipoles are thought to be important for the interactions between alpha-helices in some globular proteins, and it is possible that it is dipole–dipole interactions that initiate myosin assembly.

An interesting aspect of the importance of the C terminus concerns the role of the 28-residue charge-repeat for assembly. This charged-residue repeat gives rise to a series of bands of alternating charge along the rod and it has been proposed that complementation of these charged zones, by an appropriate stagger of the myosin rods, is a major determinant of myosin assembly (Parry, 1981; McLachlan and Karn, 1982, 1983; McLachlan, 1984). This repeat extends throughout the rod portion of myosin and, at least in LMM, is little altered between the N- and C-terminal ends of the molecule. However, despite the presence of the repeat throughout the length of the LMM molecule, our results show that it is only a comparatively small length of the sequence at the C terminus that is vital for assembly. This is not what would be expected if the complementation of the zones of alternating charge were providing the driving force for myosin assembly, because the same degree of charge complementation should be found with deletions from either terminus. Therefore, our results strongly suggest that the 28-residue charge repeat does not appear to be primarily responsible for myosin assembly.

We are most grateful for helpful advice, comments and criticisms from our colleagues in Cambridge, in particular Rob

Moir, Roy Quinlan, Reinhard Rachel, Simon Clarke, Sue Whytock and Nigel Unwin. We also thank Ross Jakes for technical assistance and Patrick Sadler for assistance with the figures. S.J.A. held a MRC *postgraduate studentship.*

References

ATKINSON, S. J. AND STEWART, M. (1991). Molecular basis of myosin assembly: coiled-coil interactions and the role of charge periodicities. *J. Cell Sci. Suppl. 14*, 7–10.

BENNETT, P. M. (1981). The structure of spindle-shaped paracrystals of light meromyosin *J. molec. Biol.* **146**, 201–221.

BRADFORD, M. (1976). A rapid and sensitive method for for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analyt. Biochem.* **72**, 248–254.

CANTINO, M. AND SQUIRE, J. (1986). Resting myosin crossbridge configuration in frog muscle thick filaments. *J. Cell Biol.* **102**, 610–618.

CARTER, P., BEDOUELLE, H., WAYE, M. AND WINTER, G. (1985). *Oligonucleotide Site-directed Mutagenesis in M13.* Anglian Biotechnology, Colchester, Essex.

COHEN, C., SZENT-GYÖRGYI, A. G. AND KENDRICK-JONES, J (1971). Paramyosin and the filaments of moluscan 'catch' muscle. I. Paramyosin: structure and assembly. *J. molec. Biol.* **56**, 223–237.

CRICK, F. H. C. (1953). The packing of α-helices: simple coiled-coils. *Acta Crystallogr.* **6**, 689–697.

CROSS, R. A. AND VANDEKERCKHOVE, J. (1986). Solubility-determining domain of smooth muscle myosin rod. *FEBS Lett.* **200**, 355–360.

CROWTHER, R. A., PADRON, R. AND CRAIG, R. (1985). Arrangement of the heads of myosin in relaxed thick filaments from tarantula muscle. *J. molec. Biol.* **184**, 429–439.

DA CRUZ E SILVA, E. F. AND COHEN, P T. W. (1987). Isolation and sequence analysis of a cDNA clone encoding the entire catalytic subunit of phosphorylase kinase. *FEBS Lett.* **220**, 36–42.

DAVIS, J. S. (1988). Assembly processes in vertebrate skeletal thick filament formation. *A. Rev. Biophys. biophys. Chem.* **17**, 217–239.

DAVIS, J. S., BUCK, J. AND GREENE, E. P (1982). The myosin dimer: an intermediate in the self-assembly of the thick filaments of vertebrate skeletal muscle. *FEBS Lett.* **140**, 293–297.

FRASER, R. D. B. AND MACRAE, T. P. (1973). *Conformation in Fibrous Proteins.* Academic Press, New York.

GAIT, M. J. (1984). An introduction to modern methods of oligonucleotide synthesis. In *Oligonucleotide Synthesis* (ed M. J. Gait), pp. 1–22. IRL Press, Oxford.

GODFREY, J. E. AND HARRINGTON, W. F. (1970*a*). Self-association in the myosin system at high ionic strength. I. Sensitivity of the interaction to pH and Ionic environment. *Biochemistry* **9**, 886–893.

GODFREY, J. E. AND HARRINGTON, W. F. (1970*b*). Self-association in the myosin system at high ionic strength. II. Evidence for the presence of a monomer-dimer equilibrium. *Biochemistry* **9**, 894–908.

HARRINGTON, W. F. AND RODGERS, M. E. (1984). Myosin. *A. Rev. Biochem.* **53**, 35–73.

HITCHCOCK-DEGREGORI, S. E. AND HEALD, R W. (1987). Altered actin and troponin binding of amino-terminal variants of chicken striated muscle α-tropomyosin expressed in *Escherichia coli. J. biol. Chem* **262**, 9730–9735.

HUXLEY, H. E. (1963). Electron microscope studies on natural and synthetic protein filaments from striated muscle. *J. molec. Biol.* **7**, 281–308.

HUXLEY, H. E. AND BROWN, W. F. (1967). The low-angle X-ray diagram of vertebrate skeletal muscle and its behaviour during contraction and rigor. *J. molec. Biol.* **30**, 383–434

IP, W. AND HEUSER, J. (1983). Direct visualization of the myosin crossbridge helices in relaxed rabbit psoas thick filaments. *J. molec. Biol.* **171**, 105–109.

KATSURA, I. AND NODA, H. (1971). Studies on the formation and physical chemical properties of synthetic myosin filaments. *J. Biochem. (Tokyo)* **69**, 219–229.

KATSURA, I. AND NODA, H. (1973). Structure and polymorphisms of light meromyosin aggregates *J. Biochem. (Tokyo)* **73**, 257–268.

KAUFMANN, E., GEISLER, N. AND WEBER, K. (1984). SDS-PAGE strongly overestimates the molecular masses of neurofilament proteins. *FEBS Lett.* **170**, 81–84.

KENDRICK-JONES, J., SZENT-GYÖRGYI, A. G. AND COHEN, C. (1971). Segments from smooth muscle myosin rod. *J molec. Biol.* **59**, 527–529.

KENSLER, R. W. AND STEWART, M. (1983). Frog skeletal muscle thick filaments are three-stranded. *J Cell Biol* **96**, 1792–1802.

KENSLER, R. W. AND STEWART, M. (1989). An ultrastructural study of crossbridge arrangement in the fish skeletal muscle thick filament. *J. Cell Sci.* **94**, 391–401.

LAEMMLI, U K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

LEINWAND, L. A., SOHN, R., FRANKEL, S. A., GOODWIN, E. B. AND MCNALLY, E. M. (1989). Bacterial expression of eukaryotic contractile proteins. *Cell Motil. Cytoskel.* **14**, 3–11.

LOWEY, S., SLAYTER, S. H., WEEDS, A. G. AND BAKER, H. (1969). Structure of the myosin molecule. I. Subfragments of myosin by enzymic digestion. *J. molec. Biol.* **42**, 1–29.

LU, R. C. AND WONG, A (1985). The amino acid sequence and stability predictions of the hinge region in myosin subfragment 2. *J. biol. Chem.* **260**, 3456–3461.

MAEDA, K., SCZAKIEL, G., HOFMANN, W., MENETRET, J-F. AND WITTINGHOFER, A. (1989). Expression of native light meromyosin in *Escherichia coli.* Observation of a powerful internal translation initiation site. *J. molec. Biol.* **205**, 269–273.

MAEDA, K., SCZAKIEL, G. AND WITTINGHOFER, A (1987). Characterization of cDNA coding for the complete light meromyosin portion of a rabbit fast skeletal muscle myosin heavy chain. *Eur. J. Biochem.* **167**, 97–102

MANIATIS, T., FRITSCH, E. F. AND SAMBROOK, J. (1982). *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

MARGOSSIAN, S. S. AND LOWEY, S. (1982). Preparation of myosin and its subfragments from rabbit skeletal muscle. *Meth. Enzym.* **85**, 55–71.

MCLACHLAN, A. D. (1984) Structural implications of the myosin amino acid sequence. *A. Rev Biophys. biophys Chem.* **13**, 167–189.

MCLACHLAN, A. D. AND KARN, J. (1982). Periodic charge distributions in the myosin rod amino acid sequence match crossbridge spacings in muscle. *Nature* **299**, 226–231.

MCLACHLAN, A. D. AND KARN, J. (1983). Periodic features of the amino acid sequence of nematode myosin rod. *J. molec. Biol.* **164**, 605–626.

MCLACHLAN, A. D. AND STEWART, M. (1975). Tropomyosin coiled-coil interactions. *J. molec. Biol.* **97**, 293–304.

MESSING, J. (1983). New M13 vectors for cloning. *Meth. Enzym.* **101**, 20–78.

MOLINA, M. I., KROPP, K. E., GULICK, J AND ROBBINS, J (1987). The sequence of an embryonic myosin heavy chain gene and isolation of its corresponding cDNA. *J biol Chem.* **262**, 6478–6488.

NAGAI, K. AND THØGERSEN, H.-C. (1987). Synthesis and sequence-specific proteolysis of hybrid proteins produced in *E. coli Meth. Enym.* **153**, 461–481.

NAKAMURA, K. AND INOUYE, M. (1982). Construction of versatile expression cloning vehicles using the lipoprotein gene of *Escherichia coli. EMBO J.* **1**, 771–775.

NOEGEL, A A., RAPP, S., LOTTSPEICH, F., SCHLEICHER, M AND STEWART, M. (1989). The *Dictyostelium* gelation factor shares a putative actin binding site with α-actinins and dystrophin and also has a rod domain containing six 100-residue motifs that appear to have a cross-beta conformation. *J. Cell Biol.* **109**, 607–618.

NYITRAY, L., MOCZ, G., SZILAGYI, L., BALINT, M , LU, R. C., WONG, A. AND GERGELY, J. (1983). The proteolytic substructure of light meromyosin. Location of a region responsible for the low ionic strength insolubility of myosin *J biol. Chem.* **258**, 13 213–13 220.

OFFER, G. (1990). Skip residues correlate with bends in the myosin tail. *J. molec. Biol.* **216**, 213–218.

O'HALLORAN, T. J., RAVID, T. J. AND SPUDICH, J. A. (1990). Expression of *Dictyostelium* myosin tail segments in *Escherichia coli*: domains required for assembly and phosphorylation. *J. Cell Biol.* **110**, 63–70.

PARKE, J., MILLER, C AND ANDERTON, B. H. (1986). Higher plant myosin heavy chain identified using a monoclonal antibody *Eur. J. Cell Biol.* **41**, 9–13.

PARRY, D. A. D. (1981). Structure of rabbit skeletal myosin. Analysis of the amino acid sequence of two fragments from the rod region. *J molec. Biol* **153**, 459–464.

QUINLAN, R. A., MOIR, R. D. AND STEWART, M (1989). Expression in *Escherichia coli* of fragments of glial fibrillary acidic protein· characterization, assembly properties and paracrystal formation. *J. Cell Sci.* **93**, 71–83.

QUINLAN, R. A. AND STEWART, M (1987). Crystalline tubes of myosin subfragment-2 showing the coiled-coil and molecular interaction geometry. *J. Cell Biol.* **105**, 403–415.

SAFER, D. AND PEPE, F. A. (1978). Axial packing in light meromyosin paracrystals. *J. molec. Biol* **136**, 343–358.

SANGER, F., NICKLEN, S. AND COULSON, A R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. natn. Acad. Sci U S.A.* **74**, 5463–5467.

SINNARD, J. M., RIMM, D. AND POLLARD, T. (1990). Identification of functional regions of the tail of *Acanthamoeba* myosin-II using recombinant fusion proteins. II. Association properties of tails with NH_2- and COOH-deletions. *J Cell Biol.* **111**, 2417–2426.

SQUIRE, J M. (1981). *The Structural Basis of Muscle Contraction.* Plenum, NY.

STEWART, M. (1975). Tropomyosin. no stagger between chains. *FEBS Lett.* **53**, 5–7.

STEWART, M. (1982). Chain register in myosin rod. *FEBS Lett.* **140**, 210–212

STEWART, M. (1988). An introduction to the computer image processing of electron micrographs of regular biological material. *J. Electron. Micros. Techn.* **9**, 301–324.

STEWART, M , BEVERIDGE, T. J. AND SPROTT, G. D (1985*a*). Crystalline order to high resolution in the shaft of *Methanosprillum hungatei*: a cross-β structure. *J. molec. Biol.* **183**, 509–516.

STEWART, M. AND EDWARDS, P. (1984) Length of myosin rod and its proteolytic fragments determined by electron microscopy. *FEBS Lett.* **168**, 75–78

STEWART, M. AND KENSLER, R. W. (1986). The arrangement of heads in relaxed thick filaments from frog skeletal muscle. *J. molec. Biol.* **192**, 831–851.

STEWART, M., KENSLER, R. W. AND LEVINE, R. J. C. (1981). Structure of *Limulus* telson muscle thick filaments *J molec. Biol.* **153**, 781–790.

STEWART, M., KENSLER, R W AND LEVINE, R. J. C. (1985*b*). Three-dimensional reconstruction of thick filaments from *Limulus* and scorpion muscle. *J. Cell Biol.* **101**, 402–411.

STEWART, M., QUINLAN, R. A. AND MOIR, R. D. (1989). Molecular interactions in paracrystals of a fragment corresponding to the α-helical coiled-coil rod portion of glial fibrillary acidic protein: evidence for an antiparallel packing of molecules and polymorphism related to intermediate filament structure. *J Cell Biol* **109**, 225–234.

SUTOH, K., SUTOH, K., KARR, T. AND HARRINGTON, W. F. (1978). Isolation and physico-chemical characterization of a high molecular weight subfragment-2 of myosin. *J. molec. Biol* **126**, 1–22.

VALENTINE, R. C., SHAPIRO, B. M. AND STADTMAN, E. R. (1968) Regulation of glutamine synthetase XII: electron microscopy of the enzyme from *E. coli*. *Biochemistry* **7**, 2143–2152.

VIBERT, P. AND CRAIG, R. (1983). Electron microscopy and image analysis of myosin filaments from scallop striated muscle. *J. molec. Biol.* **165**, 303–320.

WARD, R. AND BENNETT, P. M. (1989). Paracrystals of myosin rod. *J. Muscle Res. Cell Motil.* **10**, 34–52.

WARWICK, H. M. AND SPUDICH, J. M. (1987). Myosin structure and function in cell motility. *A. Rev. Cell Biol.* **3**, 379–421.

YAGI, N. AND OFFER, G. (1981). X-ray diffraction and electron microscopy of a light meromyosin tactoid. *J. molec. Biol.* **151**, 467–490.

YOUNG, M., KING, M V., O'HARA, D S. AND MOLBERG, P. J. (1972). Studies on the structure and assembly pattern of the light meromyosin section of the myosin rod. *Cold Spring Harbor Symp. quant. Biol.* **37**, 65–76.

(Received 1 March 1991 – Accepted 16 May 1991)